

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2001, 09:16:10 ; Search time 2192.46 Seconds
(without alignments)
23115.284 Million cell updates/sec

Title: US-08-482-402A-2

Perfect score: 3072

Sequence: 1 gaggaattgagcgccat.....atagctgcaaaaaaaaaa 3072

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_om.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pl.*

25: em_ro.*

26: em_sts.*

27: em_sy.*

28: em_un.*

29: em_vi.*

30: em_htgo_hum.*

31: em_htgo_inv.*

32: em_htgo_rod.*

33: em_htg_hum.*

34: em_htg_inv.*

35: em_htg_rod.*

36: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

Result No.	Score	Match	Length	DB	ID	Description
1	3060	99.6	3060	9	HUMTPOC	M17755 Homo sapien
2	3004	97.8	3048	6	AR092419	AR092419 Sequence
3	3004	97.8	3048	9	HUMTPOA	J02969 Human thyro
4	2991.4	97.4	3027	9	HSTPO	Y04006 Human mRNA
5	2803	91.2	2847	6	A92132	A92132 Sequence 1
6	2652	86.3	2877	9	HUMTPOB	J02970 Human thyro
7	2649.6	86.2	2845	9	HSATPO2	X17358 Human mRNA
8	2511.4	81.8	2546	6	E23825	E23825 Antigen for
9	1655.2	53.9	3237	10	RNTPO	X17396 Rat mRNA fo
10	1628.6	53.0	3291	10	MMTPEP	X07073 Mus muscula
11	1441.8	46.9	2777	10	RATTPOFR	M31655 Rat thyroid
12	970	31.6	1046	9	HUMTPOD	M55702 Human thyro
13	516	16.8	1996	9	HUMTPO07	M25707 Human thyro
14	516	16.8	160674	2	AC060811	AC060811 Homo sapi
15	510.4	16.6	1142	4	SSTPOR1	X04645 Porcine mRN
16	469.8	15.3	2238	9	S56200	S56200 myeloperoxyl
17	469.8	15.3	2261	6	A08802	A08802 H.sapiens g
18	469.8	15.3	2568	9	HUMMPV	J02694 Human myelo
19	469.8	15.3	3213	9	HSMPP	X04876 Human mRNA
20	469.8	15.3	3215	6	E12629	E12629 CDNA encodi
21	469.8	15.3	3215	9	HUMMPO	M19507 Human myelo
22	461.4	15.0	2666	10	MUSEPO	L77979 Mus muscula
23	461.4	15.0	2694	10	MUSEPP	D78353 Mouse mRNA
24	445	14.5	2740	10	MMMPO	X15313 Mouse MPO m
25	438.2	14.3	2558	9	HSEP	X14346 Human mRNA
26	433.6	14.1	2915	3	AB022196	AB022196 Clona int
27	427.8	13.9	3846	5	AF349034	AF349034 Danio rer
28	417	13.6	5510	9	D86983	D86983 Human mRNA
29	417	13.6	6847	9	AF200348	AF200348 Homo sapi
30	411.8	13.4	2814	5	AF378824	AF378824 Danio rer
31	397.4	12.9	2768	9	HSU39573	U39573 Human saliv
32	397	12.9	3179	3	AB022197	AB022197 Halocynth
33	395.2	12.9	3449	3	AB028841	AB028841 Branchios
34	386	12.6	2710	4	BOVLPO	M58150 Bovine lact
35	379.6	12.4	2780	4	AF027970	AF027970 Ovis arie
36	368.8	12.0	2634	4	CDR131675	AJ131675 Camelus d
37	344.8	11.2	172445	9	AC009471	AC009471 Homo sapi
38	338.4	11.0	435	9	S46243S1	S46243 thyroid per
39	317	10.3	300829	3	AE003475	AE003475 Drosophil
40	305	9.9	80254	2	AC019531	AC019531 Drosophil
41	301.8	9.8	2512	5	XLU68724	U68724 Xenopus lae
42	295	9.6	4871	3	DMU11052	U11052 Drosophila
43	289.4	9.4	1399	9	HUMLPO	M58151 Human lacto
44	265	8.6	1328	9	HUMTPO08	M25708 Human thyro
45	261.2	8.5	1448	5	AF364820	AF364820 Xenopus 1

ALIGNMENTS

RESULT 1

HUMTPOC

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

HUMTPOC 3060 bp mRNA
Homo sapiens thyroid peroxidase (TPO) mRNA, complete cds.
M17755
GI:4680720

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 732 to 1573)

Seto, P., Hiray, H., Magnusson, R.P., Gestautas, J., Portmann, L.,

Degroot, L.J. and Rapoport, B.

Isolation of a complementary DNA clone for thyroid microsomal

antigen. Homology with the gene for thyroid peroxidase

J. Clin. Invest. 80 (4), 1205-1208 (1987)

88008367

2 (bases 1 to 3060)

Magnusson, R.P., Chazenbalk, G.D., Gestautas, J., Seto, P., Filetti, S.,

Db 421 TGC AAAAATGCTGGATGCTCCCTTACATGCTCCCCCAAAATGCCAAAACATTTGGCT 480
Qy 525 ggcgaacaaatcacaggccatcacagagctgtcaaacacagagacacccacagatggg 584
Db 481 GCGCAACAAATACAGGCCATCACAGAGCTTGCACACACAGACACCCACAGATGGG 540
Qy 585 cgcctcaaacacggccctggcacgatggctccctccagctctatagagagcgcttcagtca 644
Db 541 CGCCTCAACACAGCCCTGGCACGATGCTCCCTCCAGTCTATGAGGACGCTTCAGTCA 600
Qy 645 gcccgaggctggaacccggctctgtatacaacagggttcccactgcccggctccggga 704
Db 601 GCCCGAGGCTGGAACCCGGCTTCTTGTACAAACGGGTTCCTCACTGCCCGCGGTCCGGGA 660
Qy 705 ggtgacaagacatgtatcaagtctcaaatgaggtgtgacagatgatgaccgtattc 764
Db 661 GGTGACAAAGACATGTCTAAGTTTCAATGAGGTTGTACAGATGATGACCGCTATTTC 720
Qy 765 tgacctctgatggcatggggacaaatatacatcgaccacagacatcgcttcacaccacagag 824
Db 721 TGACCTCTGATGTCATGGGACAAATACATCGACACAGATCGGTTACACCCACACAG 780
Qy 825 caccagaaagctccttcggggaggggtctgactgcccagatgactgtgagaaacaaaa 884
Db 781 CACAGCAAAAGCTGCTTCGGGGAGGGGCTGACTGCGCAGATGACTTGTGAGAACCAAAA 840
Qy 885 cccatgtttcccatatacaactcccgagagagcccgcccgcccgagcccgctgtct 944
Db 841 CCCATGTTTTCCCAATACAACTCCCGAGAGGCGCCCGCGCGGCGACCGCTGTCT 900
Qy 945 gccctttacgcctcttcggccgctcgggcaccggggaccaaaggcgctctttggaa 1004
Db 901 GCCCTTACCGCTCTTCGGCGCGCTCGCGCACCGGGACCAAGCGCGCTCTTTGGGAA 960
Qy 1005 cctgtcaacggcgaacccgagcgagagatgaacgggttgacctgttccctggacgctc 1064
Db 961 CCTGTCCAGGCCCAACCCCGCGGACGACAGATGAACGGGTGACCTCGTTCTTGAGCGCTC 1020
Qy 1065 caccgttatggagctcccgcccttagagagcagctcggaactggaactggaactggccga 1124
Db 1021 CACCGTGTATGGCAGCTCCCGCGCCCTTAGAGAGGACGTGCGGAACCTGGACAGTGGCCGA 1080
Qy 1125 agggctgtccgctccacggccctccggactccggcgcgcgcgcctaccctgaccttct 1184
Db 1081 AGGGCTGTCCGCTCCACGCGCGCTCCGGGACTCCGGCGCGGCTACCTGCTGCTTCT 1140
Qy 1185 gccgcaacgcgctcgctggcctgtgcccagcccgagcccggaaccccgagagacccg 1244
Db 1141 GCCGCCACGCGCGCTGCGCGCTGTGGCGCGAGCCCGGCATCCCGGAGACCCGCGG 1200
Qy 1245 gccctgtcttcggcgagagcgccgcccagcgaggtccctcccttcagcggaactgca 1304
Db 1201 GCCCTGTCTTCGCGGAGACGCGCGCGCAGGAGTCCCTCCCTCCCTCCCTCCCTCC 1260
Qy 1305 caecgtgtggctgtgcgagacaaacccgctggcgcgcgctcaaggccctcaatggcga 1364
Db 1261 CACCGTGTGGCTGCGGAGCACACACCGGCTGCGCGCGGCTCAAGGCCCTCAATGCCGA 1320
Qy 1365 ctggagcgcgacccgctgtaccagggagcgcgcaaggtctggcgctctgacacagat 1424
Db 1321 CTGAGCGCGGACGCGCTGTACCAGGAGCGCGCAAGGTCTGGCGCTCTGCAACAGAT 1380
Qy 1425 catcaacctgaggatatacatcccaagatcctgggaccccgagccttccagcgagtcgt 1484
Db 1381 CATCACCTTGAGGATTAATATCCCAAGGATCTCGGACCCGAGGCTTCCAGCATAGCT 1440
Qy 1485 gggctccatgaagcttatgactcacgcgaaccccaactgtgccacgtgtctccac 1544
Db 1441 GGGTCCCTATGAAGGCTATGACTCCACGCCCAACCCCACTGTGTCCAAACGTTCTCCAC 1500
Qy 1545 agccgcttccgcttcggccatgccaagatccacccgctggtagagggctggagcgccag 1604

Db 1501 AGCCCGCTTCCTCGCTTCGGGCCATGCCACGATGCCACCCGCTGGTGAGGAGGTGACGCCAG 1560
Qy 1605 cttccaggagcaccggcagctcccgggctgtgctgcaccagcgttctcttcagcccatg 1664
Db 1561 CTTCCAGGAGCACCCGACCTGCCGGGCTGTGGCTGCACACAGCTTCTTTCAGCCCATG 1620
Qy 1665 gacattactccgtggaggttggtttggaccactaaatcagaggcctcttctgaaagacagc 1724
Db 1621 GACATTACTCCGTGGAGTGGTTGGACCCACTAATACGAGGCTTCTTCTGCAAGACGACG 1680
Qy 1725 caaatgcaaggtgcaggatcagctgatacagagagctgcaggaagagctcttctgtct 1784
Db 1681 CAAATGCAAGTGCAGGATCAGCTGATGAACGAGGAGCTGACGGAAGGCTCTTTGTGCT 1740
Qy 1785 gcccaattccagcacccttgatctggctccatcaacctgcagagggcgcgagcacagc 1844
Db 1741 GTCCAAATTCACGACCTTGAATCTGGCTGCCATCAACCTGCAGAGGGCGGGACACCG 1800
Qy 1845 gctgccaggttacaatgagtgaggagttctgctggcctgcctgcctggagacccccgc 1904
Db 1801 GCTGCCAGGTTAAATGAGTGGAGGAGTTCGTGCGGCTGCCTCGCTGGAGACCCCGC 1860
Qy 1905 tgaactgagcagacccatcgccagcagggagctggccgacaaagatcctggactgtacaa 1964
Db 1861 TGACCTGACACAGCCATCGCCAGCAGAGCGGTGGCCGACAAAGATCTGGACTGTACAA 1920
Qy 1965 gcatcctgacaacatcgatgtctgctggagccttagctgaaacctctcccccagggc 2024
Db 1921 GCATCTGACAAATCGATGTCTGGCTGGAGGCTTGTAGTGAATACTCTCCTCCAGGGC 1980
Qy 2025 tcggcagggccctgtttgctgtctcatctgggaagcagatgaaaggctctctgcggagc 2084
Db 1981 TCGGACAGGGCCCTGTTTGCCTGTCTCATTTGGAAAGCAGATGAAGGCTCTGCGGGATG 2040
Qy 2085 tgaactggttttgaggagaaacagccactcttcacggatgcacagagcgctgagctgga 2144
Db 2041 TGACTGGTTTTGGTGGGAGAACAGCCACCTTTCACGATGTCACAGAGCGGTGAGTGA 2100
Qy 2145 gaagcactcctgtctcggtcatctgtgacaacactggcctcaccaggtgccccagga 2204
Db 2101 GAAGCACTCCCTGTCTCGGGTCACTGTGACAACTATGGCCTCACCAGGGTGCCCATGGA 2160
Qy 2205 tgccttccaagtccggcaaatcccccgaacttttgagctgtgtgacagcatcactggcat 2264
Db 2161 TGCTTCCAAAGTCGGCAAAATTCCTGAGACTTTTGTAGTCTTGTGACAGCATCCTTGSCAT 2220
Qy 2265 gaactggaggcctggaggaaaccttctcctaagacagacaaagtgtgcttcccagagag 2324
Db 2221 GAACCTGGAGGCTGGAGGGAAACCTTCTCTCAAGACGACAAAGTGTGGCTTCCAGAGAG 2280
Qy 2325 cgtgagaatgggactttgtcacctgtgagagctctggagcgctgctgtgtattc 2384
Db 2281 CATGGAGAAATGGGACTTTGTGCACCTGTGAGGAGTCTGGAGGGCGGTGCTGTGTATTTC 2340
Qy 2385 ctgcccggcacgggtatgagctccaaaggccggagcagctcacttgaccacaggaagatg 2444
Db 2341 CTGCGGACACGGGTATGAGCTCCAAGCCGGGAGCAGCTCACTTGACCCAGGAAGATG 2400
Qy 2445 ggaattccagcctccctctgcaaatgtgaacagagtgcagagcggtgccccaccccc 2504
Db 2401 GGATTTCCAGCCTCCCTCTCTGAAAAGATGTGAACAGAGTGTGACAGCGTGTGCCACCCCC 2460
Qy 2505 ctgcccagcctctggagggtgcagaaaaacccaaaggcgcttccagtgctctctgcgcga 2564
Db 2461 CTGCCACGCTCTGGCGAGGTGCAGAAACACCAAAAGCGCTTCCAGTGTCTCTGCGCGGA 2520
Qy 2565 cccctacaggttagagacgatggagagaacctgcgtagactccggagggctcctcgggt 2624
Db 2521 CCCCTACGAGTTAGGAGACGATGGGAGAACCTCCGTAGACTCCGGAGGCTCCTCTCGGC 2580
Qy 2625 gacttggatctcccatcgctgctgctctgtgacggaggtctgcaggtctcactc 2684
Db 2581 GACTTGGATCTCCATGTCTGCTGCTGCTGATCGGAGGCTTCGAGGCTCTCACCTC 2640

RDHPRWASNTALARWLPPVYEDGFSQPRGNPGLYNGFLPPLPVREVTRHVIOVNSNE
 VTDDRYDLMLMANGQYIDHDIATPQTSKAAFGGGADCOMTCENQPCPFIQJLPE
 EARPAGTACALPFISSAACGTGDOGALFNLSTANPROOMNGLTSLFLDSTVYGSSP
 ALEROLRNWTSAGELLRLVHARLRDSGRAYLPVPPRRPAACAPERGIPEGTRGPCFLA
 GDGRASEVPSLTALHTLWLREHNLAAALKALNAHWSADAVYQEARVYVGHALHQTIL
 RDIYIPRIEPAFQIQVGYEYDSTANPTVSNVSTAFRFGHTLHPLVRLDASF
 QEPHLPGLWLQWAFSPWTLRLRGYNREFFGLQPRLETAPDLSTAIASRSYADKIL
 DLYKHPDNIMVGLGLAEFLPARTGPIFLAGLIKQKALRDGDFWFWNSHVFDIA
 ORRELEKHSLSRVICDNLTRVPMDAFOVGFPEFESCDSTIPGNMLCAWRETPFOD
 DKCGFPESEVNDCHTCESGRVLVYSCRHGYELOGRQSLCTOGWDFQPLCKDV
 NECADGAPPCBASACRNTKGGFQCLCADPYELGDGRTVCVDSGRLPRATWISMSLA
 ALLIGFAGLTSTVICRWTRTKSLPLISETGGGTPELRCRKHQAVGTSPORAAAQD
 SQESAGMEGRDTHRLPRAL"

BASE COUNT 632 a 883 c 829 g 533 t
 ORIGIN Chromosome 2pter-q11.

Query Match 86.3%; Score 2652; DB 9; Length 2877;
 Best Local Similarity 93.7%; Pred. No. 0;
 Matches 2857; Conservative 0; Mismatches 15; Indels 176; Gaps 2;

QY 18 catttcagaagattacagccgtgaaattactcagcagtgcaagtggtgctgagaagagga 77
 Db 1 CATTTCAGAAAGATTACAGCCGTGAAATTTACTCAGCAGTGCAGTTGGCTTGAGAAAGGA 60
 QY 78 aaaa-----agaatgagagcgtggctgtgctgtctgtcagcgtggttatggcctgcaca 132
 Db 61 AAAAAGGTGAGATGAGAGCGCTCGTGTGCTGTCTGTCTACAGTGTGTTATGGCCGTGCACA 120
 QY 133 gaagccttctcccttcattctcagagagggaagaactcctttggggaagcctgagag 192
 Db 121 GAAGCCTTCTTCCCTTCATCTCAGAGAGGAAAGAACTCCTTTGGGAAAGCCTGAGGAG 180
 QY 193 tctcgtgtctctagctgttgggaaagcaagcgcgtgtgtgagacacgcctatgtaagcc 252
 Db 181 TCTCGTGTCTTAGCTCTTGGAGGAAGCAAGCGCTGTGTGGACACCGCCATGTAGGCC 240
 QY 253 acgatcagagaaacctcaagaaagaggaactcttcttgagtcagctctctgtctttt 312
 Db 241 ACATGACAGAGAACCCTCAAGAAAGAGGAATCCTTTCTCCAGCTCAGCTTCCTCTTTT 300
 QY 313 tccaaactctctgagccaaacagcgagtgattgcccagcagcagagataatggaaaca 372
 Db 301 TCCAAACTTCTCAGCCCAACAAGCGGAGTGATTGCCCGAGCAGCAGAGATAATGGAAACA 360
 QY 373 tcaatacagcagatgaaagaaagtaactcaactgaaactcaacaaactcaacagatccaag 432
 Db 361 TCAATACAAGCGATGAAAAGAAAAGTCAACCTGAAAACCTCAACAATCACAGCATCCAAGC 420
 QY 433 gatgctttatcagaagatctgtgagcatcttgcaaacatgctctgagatctccctttac 492
 Db 421 GATGCTTTATCAGAAGATGCTGTGAGCATCTTGTCAAAACATGTCTGGATGTCTCCCTTAC 480
 QY 493 atgctgcccccaaatgccccaaacttgcctggcgaacaaatacagcggccatcacagga 552
 Db 481 ATGCTGCCCCCAAAATGCCAACACTTGCTGGCGAACAAATACAGGCCCATCACAGGA 540
 QY 553 gcttgcacaacacagacacccacagatggggcgctcccaacacgcccctggcaacgatg 612
 Db 541 GCTTGTCAACAACAAGACACACCCAGATGGGGCGCTCCCAACACGGCCCTGGCACGATGG 600
 QY 613 ctccctccagctctatgagacggtctcagtcagcccgaggtggaaacccggtctttg 672
 Db 601 CTCCCTCCAGCTATGAGGACGGCTTCAGTACGCCCGGAGGCTGGAACCCCGGCTTCTTG 660
 QY 673 tacaacgggttccactgcccgggtccggagggtgacagacatgctcattcaagttca 732
 Db 661 TACAACGGGTTCCTCCACTGCCCCCGGTCCGGGAGGTGACAAGACATGTCATTCAAGTTTCA 720
 QY 733 aatgaggtgtcacagatgatgaccctattctgacctctctgatggcatgggacaatac 792
 Db 721 AATGAGGTTGTACAGATGATGACCGCTATTCTACCTCTCTGATGGCATGGGACAATAC 780

QY 793 atcgaaccacgacatogcgttccacaccacagacagaccagcaaaagtccttcggggaggg 852
 Db 781 ATCGAACACGACGATCGCGTTTACACACACAGACGACCAAGAGTGCCTTCGGGGAGGG 840
 QY 853 tctgaactcgaagatgactgtgagaacaaacccatgttttcccatatacaactcccgag 912
 Db 841 GCTGACTGCCAGATGACTTGTGAGAACCAAAACCCATGTTTCCCATACAATCCCGGAG 900
 QY 913 gagcccgccgcccgcgggacgcgctgtctgacctctctacgcctcttcggccgctcgc 972
 Db 901 CAGCCCGCCCGCGCGCGGGGACACCGCCTGTCTGCGCTTCTACCGCTCTCTCGCGCGCTGC 960
 QY 973 ggcaccggggaccgaagcgcgctcttttgggaacctgtccacggcccaaccccgaggcagag 1032
 Db 961 GGCACCGGGGACCAAGGCGCGCTCTTTGGAAACCTGTCCACGGCAACCCGCGGAGCAG 1020
 QY 1033 atgaacgggttgacctgttcttggaacgcgtccacccgtgtatgagcagctcccgcccta 1092
 Db 1021 ATGAACGGGTTGACCTCGTCTTCTGGACGCGTCCACCGTGTATGGCAGCTCCCGCGCCCTA 1080
 QY 1093 gagagcagctgcggaactgacacagtgccgaagggctgcctccgctccacgcgcgcctc 1152
 Db 1081 GAGAGGAGCTGCGGAAGTGGACCAAGTCCGGAAGGCTGTCTCCGCTCCACGGCGCCTC 1140
 QY 1153 cgggactccggccgcgctacactgcctcttctgcccgaacgcgcgctcgcgctgtgog 1212
 Db 1141 CGGACTCCCGCGCGCTACTGCTCCCTTCGTGCGCGCACGCGGCGCTCGGCTGTGCG 1200
 QY 1213 cccgagcccggaaccccgagagaccgcggccctgcttctcctggcgagacggccgc 1272
 Db 1201 CCCGAGCCCGCATCCCGGAGAGACCGCGGGCGCTGTCTTCTGGCGGAGACGCGCGC 1260
 QY 1273 gccagcaggtccctccctcctcaacgcactgcacacgctgtgctgcgagcacacccg 1332
 Db 1261 GCCAGCGAGGTCCCTCCCTCCTGACGGCACTGACACGCTGTGGCTGCGCGAGACAAACCG 1320
 QY 1333 ctggcgcgcgctcaagccctcaatgcgcactggagcgcgagcgctgtaccagag 1392
 Db 1321 CTGGCCGCGCGCTCAAGGCCCTCAATGGCACTGGAGCGCGAGCGCTGTACAGGAG 1380
 QY 1393 gcgcgaaggtctgtggcgctctgcacagatca tcaacctgagggattacatccccag 1452
 Db 1381 CGCGCAAGGTCGTGGGCGCTGTGCACCAAGATCATACCCCTGAGGAGATTACATCCACG 1440
 QY 1453 atcctgggacccagcgcctccagcagctacgtgggtccctatgaagctatactccacc 1512
 Db 1441 ATCTTGGACCCCGAGGCGCTTCCAGCAGTACGTGGGTGCCCTATGAAGGCTATGACTCCAC 1500
 QY 1513 gccaaccccaactgtgtccaaagctgttctccacagcgcttccgcttcgcccagtcacag 1572
 Db 1501 GCCAACCCCACTGTGTCCAACGTTCTCCACAGCGCGCTTCCGCTTCGGCCATGCCAG 1560
 QY 1573 atcacccgctggtgagaggtggaacccagcttccaggagcacccccacactgccccgg 1632
 Db 1561 ATCCACCCGCTGTTGAGGAGGCTGGACGCGAGTCTCCAGAGACACCCCGACCTGCCCGG 1620
 QY 1633 ctgtggtgcaccaggttcttccagcccatggaacattactcgtggaggtggtttggac 1692
 Db 1621 CTGTGGCTGACCAAGGCTTCTTTCAGCCCATGGACATTAATCTCCCTGG----- 1667
 QY 1693 ccactaaacagggcctctgtgcaagaccagccaaactgcaggtgcaggtcagctgatg 1752
 Db 1668 ----- 1667
 QY 1753 aacgagagctgacggaaggtctcttgtgtccaatccacagcaccttggatctggog 1812
 Db 1668 ----- 1667
 QY 1813 tccatcaacctgcagagggcgccggaccacacggctgcagaggttacaatagtgaggaggg 1872
 Db 1668 -----AGTTTACAATGAGTGGAGGAG 1689
 QY 1873 ttctgcgctcgtcgtcgtggagaccccgctgacctgagcacagcctcagcagcag 1932

FRTL5 rat thyroid cells
JOURNAL Nucleic Acids Res. 17 (20), 8380 (1989)
MEDLINE 90045972
FEATURES
 source Location/Qualifiers
 1..3237
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /cell_line="FRTL5."
 42..2786
 /notes="thyroid peroxidase (AA 1-914)"
 /codon_start=1
 /protein_id="CAA35257.1"
 /db_xref="GI:57383"
 /db_xref="SWISS-PROT:P14650"
 /translation="MRTLGAMVYLVVMGTAFILPFLRRSDILGKMTSHVSVIVE
TSOLLVDNAVYTMKRLKRRKLVSLFSKLPSTSGAISRAAEIMETSIOVMK
REOSQFSTDAISLILATLANGLPMLPRCPDPTCLANKYRPIPTGVGNRRDHPW
GASNTALAWLPVYEDGESQPRGNPNFLYHGFLPPVREVRHLLOVSEATEDD
QYSDFLPVWGQYIDHIALFTPOSTAAAFWGGVDCQLTCENQNPFCFIQLPSNSRRT
WSSAGLLRVNTLHLSGRAYLPFAACAPEGAPHANRTPCFLMGDRASEVPALA
AVHTLWLRNHLRTAFKINTHWSANTAYOEARKVVGALHQIITMDRYIPKILGPA
FROYGVPYENPTVNPVTSNVSFAAFREGHATVPLVRLNLTDFODHTLPLQLH
DVEFRWRLIOEGGLDPIVRGLLARPALQVQLNNEELTERFLVLSNVGLDLASL
NLORGHDGLPGTNEWREFCGLSLDTGALNKAIAINRSWVIMELYKHADNDVWL
GLAEKFLPGARTGPLFACILIGKOMKALRDGDFWENSHVFTDAQROEKEKSLPRV
ICDNTGLTRVPDAFRIGFPQDFESCEIIPSDMLRLWRETFQDDKRVPEKVDGN
FVKCESKGLVLYSCFHYKLOGEOVTCQWDSPEPVCKDVNCEADLTTHPPCHS
SACNTKSGSFQVCVCTDPMGLGDEKTCIDSGRLPRASWSVIALGALLJGLLSLW
VICRWTHADKKSLILITERTVMSGFRKSOESGISPKAEVQDAEQEPAYGSRVLLCE
" *
BASE COUNT 816 a 874 c 831 g 716 t
ORIGIN

Query Match 53.9%; Score 1655.2; DB 10; Length 3237;
Best Local Similarity 76.4%; Pred. No. 1.4e-292;
Matches 2105; Conservative 0; Mismatches 613; Indels 42; Gaps 4;

Qy 82 agaagagagcggtgctgtctgtcaacgctgtgttgctgacctgacagagacgttc 141
Db 39 AGAATGAGACACTTGGAGCTATGGCAGTAATGCTGGTTGTATGGGAACTGCAATTTTC 98

Qy 142 ttccctctatctgagagaaagaactcctttgggaaagcctgaggagctctgtgctc 201
Db 99 CTCCTTTCTCCTGAGAAGCAGACATCTTGGGGGGAAGACTATGACATCCCATGTT 158

Qy 202 tctagcgtcttgagaaagcaagcgcctggtggaacccgcctatgtaacgcaatgag 261
Db 159 ATCAGTGTGTAGAAACGAGCAGCTTCTGGTGGGACATGACAGTCTACACACCATGAAA 218

Qy 262 agaaacctcaagaaagagaactcttcttgagctcagctgtgtcttttttcaaaactt 321
Db 219 AGAAACCTCAAGAAAGGGAGTCTTCTCCAGCCGAGCTTCTCTTCTTCCAAAGCTG 278

Qy 322 cctgagcaacaagcggagtgttggccgagcagcagagataatggaaacatacaataca 381
Db 279 CCAGAGTCCACAGTGGAGCTATTTCGGAGCAGCAGAGATCATGGAACATCAATACAA 338

Qy 382 gcgatgaaagaaagtcaacctgaaactcaacaatacacagcatccaacggatgcttta 441
Db 339 GTCATGAAACG-----TGAACAATCACAGTTCTCCACGGATGCACCTA 380

Qy 442 tcagaaagtctgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 501
Db 381 TCACAGACATCTCTGGCCACAAATTCGCAACCTGTCCAGAGTCTTGCCTTTTCATGCTGCA 440

Qy 502 ccaaaatgcccaacacftgcctgagcaacaataacagggcccatcacagagagcgttgeac 561
Db 441 CCGAGATGCTTGACACTGCTGCGCACAAATTCGCAACCTGTCCAGAGTCTTGCCTTTTCATGCTGCA 500

Qy 562 aacagagaccacccagatggggcgctcccaacagcgcctggcacgatggctccctcca 621
Db 1563 CACAGCTCTTCTTCAGGCGCTTGAGGCTTATCCAGGAGAGGTGGTTGGATCCAAATAGTA 1622

Db 501 AACAGAGATCACCCAGATGGGAGGCTCCAAACACAGCCCTAGCAAGATGGCTGCCTCCT 560
Qy 622 gtctatgagagcggttcagtcagccccagcgtggaaccccggtcttcttctgtacaacggg 681
Db 561 GTCTACGAAGATGGTTTTCAGTCAGCCAGAGCTGGAACCCCTAATTTCTTATACCATGGC 620
Qy 682 ttcccaactgcggcggtccgggaggtgacaaacatgtcattcaagttaaaatgaggtt 741
Db 621 TTCCCACTGCCCGGTAGCGGAGGTGACAAGACACCTCATTTCAAGTTCAAAATGAGGCT 680
Qy 742 gtccagatgatgacgcctattctgaactcctctgagcgtgagcagcagcagcagcagc 801
Db 681 GAGATGAAGATGACACAGTACTCTGATTTTCCCGGTGGGAGCAGTACATCGATCAT 740
Qy 802 gacatcggttcacaccacagacagcagcagcagcagcagcagcagcagcagcagcagc 861
Db 741 GACATTGCTCTCACACCACAGACACACAGACGCTTCTCCCATACAGTTCCTCCAACTCCTCA 800
Qy 862 agatgactgtgagaacccaaacccatgttttcccatcaaaactccccgagggagccgg 921
Db 801 CAGTTGACCTGTGAGAACCAAAACCCCTGCTTCCCATACAGTTCCTCCAACTCCTCA 860
Qy 922 ccggcgcgcgccacgcctgtctgcccctctaccgctcttccgcccctgcgcagcagcgg 981
Db 861 CGGAC-----CACTGCATGCTTACCTTTTACCGCTCTCTCAGCCGCTGTGGCAGCTGGG 914
Qy 982 gaccaaggcgcgctctttggaaacctgtccacggccaaacccagagcagcagcagcagcgg 1041
Db 915 GACCAAGGAGCACTCTTTGGCAACCTGTCTGCAGCCAAATCCGAGGAGCAGATGAATGGC 974
Qy 1042 ttgacctgcttctggaagcgtccacgctgtatggcagctcccccgccctagagagcag 1101
Db 975 TTGACCTCTCTTCTGTATGCTTCCACGGTGTATGGCAGCTCCCGAGGTGTGAGAAGCAG 1034
Qy 1102 ctgcggaactgaccagtgccgaagggtgctccggctgctccagggcgccctccgggactcc 1161
Db 1035 TTGCGCAACTGGAGAGCGTCTGCAGGGCTGCTGCGTGTCAACACCTCTCCACCTAGACTCT 1094
Qy 1162 ggcgcgcctactgcctctgtgcccgcacgcgcgcctgcgcgcctgtgcgcgcgcgcgc 1221
Db 1095 GGCGGTGCTTACTTGCCTT-----CGCATAGCCGCTGTGCTCCAGAGCCT 1142
Qy 1222 ggcaaacccgggagagacccgcggccctcttcttggcggagacggcgccgcgcgcag 1281
Db 1143 GGTGTCTCACACGCCAACCGCAGCCTGCTTCTTCTGGTGGTGTAGTGCCTGCGCAGCGAG 1202
Qy 1282 gtccctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1341
Db 1203 GTCCCGCCCTGGCAGCAGTGCACACCTTGTGGCTGCGGAGCACAAACCGCTGGGTACA 1262
Qy 1342 gcgctcaaggccctcaatgctgcactggagcgcgcgcgcgcgcgcgcgcgcgcgcgcag 1401
Db 1263 GCTTTCAAGCCATCAACACATTTGGAGCGCCAAACACTTGCCTTACCAGGAGGACCGCAAG 1322
Qy 1402 gtccggcgctctgcaccagatcatccctcagggaggtatcaatccccaggatccctggga 1461
Db 1323 GTGGTAGGGGATTCACACAGATCATCACCATGAGGAGTATATATCCCAAGATCCTGGGT 1382
Qy 1462 cccgagggccttccagcagtcagtggttccctatgaagggtatgactccaccgcgaacccc 1521
Db 1383 CCCGATGCTTTCAGGAGTATGTGGGCCCTATGAAGGCTATAAACCCACCGGTGAACCT 1442
Qy 1522 actgtgtcaacgcttctccacagcgccttccgcttccgcttccgcttccagcagtcaccc 1581
Db 1443 ACTGTGTCCAAAGCTCTCTCTACTGCTGTCTTCCGCTTTCGCTTTCGCTTTCGCTTTC 1502
Qy 1582 ctgtgtgagagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1641
Db 1503 CTGGTAGAGCAGTAAACAGGACATTCAGGAGCACCACAGAGCTCCCGAGGCTGCGAGCTG 1562
Qy 1642 caccagggcttctcagcccatggacattactcctggtggaggtggttggagccactaata 1701
Db 1563 CACAGCTCTTCTTCAGGCGCTTGAGGCTTATCCAGGAGAGGTGGTTGGATCCAAATAGTA 1622

Qy	1702	cgaggcctcttgcagaccagccaaactgcaggtgcaggatcagctgataaacaggag	1761
Db	1623	AGAGGCTCTCTGCGAAGACACGCCAAGCTGCAAGTACAGGAGCAACTGATGAATGAGGAA	1682
Qy	1762	ctacaggaagcctcttctgtctccaatccagcaaccttggatctggcgtccatcaac	1821
Db	1683	CTGACCGAGAGGCTCTTTGTGTGTCTAATGTGGGCACCTTTGGATCTGGCATCAGTGAAC	1742
Qy	1822	ctcacagggccggagacacggctgcaggttacaatgattgagagggaggttctcgcc	1881
Db	1743	TTGCACAGGGGCGGACCAACGGCTTACAGGCTACAATGAGTGGAGAGAGTTCTGTGGC	1802
Qy	1882	ctgcctcgcctggagaccccgctgacctgagcacagccatcgccagcagagcgtggcc	1941
Db	1803	TTGTACGGCTGGACACAGGAGCTGAGCTGACACAGGCCATTGCCACACAGCATGCTC	1862
Qy	1942	gacaaatcctggactgttacaagcatcctgacaacatcgatgtctggctgggagcgcta	2001
Db	1863	AACAAGATAATGGAGTTATACAAGCATGCTGACAACTTGATGTCTGTTGGGAGGCGCTG	1922
Qy	2002	gctgaaaactctctcccagggtcctggacagggccccctgtttgacctgtctcatattgggaag	2061
Db	1923	GCTGAAAAGTTCTTACCGGAGCCGCGCACCGGCTCTCTGTTTGCATGTATCATTTGGGAAG	1982
Qy	2062	caataaagcctcgcggagcgtgaactggtttgggtgggagaacaccacgcctctcaag	2121
Db	1983	CAGATGAAGGCTCTGAGGGATGGGACAGGTTTGGTGGGAGAACGCCATGTCTTCACA	2042
Qy	2122	gatgcacagagcgtgagctggagagcactccctgtctcgggtcatctgtgacaacact	2181
Db	2043	GAGCTTCAGAGCGAGNACTGGAAAAGCAATTCATACCTCCGGTCTACTCTGTGACACACC	2102
Qy	2182	ggcctcacagggtgccatggatgccttccaagtgcggcaattccccgaagactttgag	2241
Db	2103	GGCCTCACAGAGTACTGTGTGATGCCTTCCGTATTGGAAAGTTCCCCAGGACTTTGAA	2162
Qy	2242	tcttgtcacagcatacctgcatgaactgcagcctggaggggaaaaccttctcaagac	2301
Db	2163	TCCGTGTAGGAAATCCCTAGCATGGACCTCAGACTGTGGAGGGAGAGCTTCCACACAGAC	2222
Qy	2302	gacaagtggcttccagagagcgtgagaatgggaattgggaactttgcactatgagaactc	2361
Db	2223	GACAAGTGTCTTCCACAGAAGGTGGACATGGGAATTTGTGCACTGTGGAAGAACTC	2282
Qy	2362	ggagggcggtgctggtgtattcctgcgcgcagcggatagagctccaaaggccggagcag	2421
Db	2283	GGGAAGCTGGTACTGGTGTATTCCTGTTTCCATGATACAGCTGCAGGCCAGGACGAG	2342
Qy	2422	ctcactgcacccaggaaggaatggatttccagctcccccctctgcaagatgtgaacgag	2481
Db	2343	GTCACTGTACCCAGAAATGGATGGGACTCAGAGCCTCCTGTCTGTAAAGATGTTAATGAG	2402
Qy	2482	tgtgcagacgggtgccaccccccttgcacgcctcgcagaggtgcagaaacacccaaggcc	2541
Db	2403	TGTGCGAGATCTGACACACCACCTTTGCCACTCTCCGCAAAAGTGCAGAACACCCAGGGA	2462
Qy	2542	ggcttccagtgtctcgcgcgaacccctacaagtttagagacgcatggagaaacctgccta	2601
Db	2463	AGCTTCCAGTGTGTGACACACCCCTTACATGCTTAGGTGAGGATGAGAAGCTGCATA	2522
Qy	2602	gactcggagggtctctcgggtgtgactggatctccatctgcgtggctcgtctgctgac	2661
Db	2523	GATTCYGGCAGGCTACTCGGGCATCTCTGGGTCTCCATTGTCATTTGGGTGCACCTCTCAT	2582
Qy	2662	ggaggctctgcaggtctcaacctgcagcgtgtattgcaggtggacacgcactggcactaaa	2721
Db	2583	GCTGGTTTGGCCAGTCTCAGCTGGACTGTAAATTTGCAGGTGCACACATGCTGATAAGAG	2642
Qy	2722	tccaaactgcocatctcggagacagcgggagaaactccccagctgagatcgcgaaagcac	2781
Db	2643	TCCACATTTGCTGATCACCAGGA-----GAGTGAACATTGGAGTCCAGGATTCAGAAAGAT	2696

Qy	2782	caggccgtagggacctcacccagcagcgccgacgcttcaggactcggacgacgaagaagtgcct	2841
Db	2697	CAGGAGAGTGGGATTTCACCAACAAGAAGCCGAGGTTCAGATGCTGAACAGAAACCGGCT	2756
RESULT 10			
MMTPR 3291 bp mRNA ROD 25-AUG-1999			
LOCUS Mus musculus mRNA for thyroid peroxidase.			
DEFINITION X60703			
VERSION X60703.1 GI:297160			
KEYWORDS peroxidase.			
SOURCE mouse.			
ORGANISM Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 3291) Kotani,T., Umeki,K., Yamamoto,I., Takeuchi,M., Takechi,S., Nakayama,T. and Ohkaki,S. TITLE Nucleotide sequence of the cDNA encoding mouse thyroid peroxidase JOURNAL Gene 123 (2), 289-290 (1993)			
MEDLINE 93154601			
REFERENCE 2 (bases 1 to 3291)			
AUTHORS Sachiya,O.			
TITLE Direct Submission			
JOURNAL Submitted (26-JUN-1991) O. Sachiya, Central Laboratory for Clinical Investigation, Medical College Hospital, Miyazaki Medical College, Kiyotake, Miyazaki 889-16, JAPAN			
FEATURES Location/Qualifiers			
source 1..3291			
/organism="Mus musculus"			
/strain="C57BL/6"			
/db_xref="taxon:10090"			
/tissue_type="thyroid"			
/cell_type="follicular"			
/clone="pmTPO-75"			
74..3281			
/partial			
/note="peroxidase"			
/evidence=experimental			
74..2818			
/EC_number="1.11.1.7"			
/codon_start=1			
/product="peroxidase"			
/protein_id="CAA43114.1"			
/db_xref="GI:4539541"			
/db_xref="MGD:MGI:98813"			
/db_xref="SWISS-PROT:P35419"			
/translation="MRTLGAMIMLVVGVTVIFLSFILSRDILCGTKMSHVSIVAVE" TSOLMDHAVNTMKRLKKREVLSPALLSFPEKLPESTSCAISRRAEIMETSIVQM REQQSFDALSADILGTIANISGLPLPMLPPPCPDCLANKYRPITGCANNRDHPRWK GASNTAKRNLPVPYEDGFSGPKWNNFIIHGFLPPPREVTNRHLLOVNEATEDDT QYSDFLFWQQOYIDHIALPOSTAAFGWGDCQLTCENQCFFIQIPLSSNGSTTT ACSPFYRSACCTGDGALFNLSAANPRQMNGLTSFLDASTVYGSSPGVEKLRNR WSSGAGLLRVTLHDAGRAYLPATAACAPECPGTPRTNRTPCFLAGDGRASEVPLA AVHTLWLREHNRLSAFKAIKNKWSANTAYOAEARKVGALHQIITMRDYIPKILGPDA FRVYPGPNLTNPVTNVSNFSTAARFHGHVPLVRRLTDFOEHTELPRLQLR DQYFVPRNLIOEGDLDPVIRGLLARALKVQGOLMNEELTERLFVLUSNVGTLDLASL NQLRGRDHGLPDYNWEREFCGSLRETPAELNKAIANSRMVNKTMDLYKHADNDIVWL GGLAEKFLPGARTGFPACIIGKQMKALRGDRFWEVENTNYFTDAQKELEKHSPLRP ICDNTGTRVVDARFGKPKPDFESCDIPSMDELWRETFPODDKCVFPEEVDNCG FVICFEESKLVLYVCFPHYKLOGEOVTCOKGWDSPPVCQDVNSCADLTHPPCHPI SAQCKNTKGFCQCVCTDPYVLGEDEKTCIDSGRLPRASWSVIALGALLIGLASLTWIC VICKRWTHADKKATLPTERTVTQTSGCRKRSGRGLSPHKAAAQDTGQPASGSRVLLCE "			
mRNA			
CDS			
BASE COUNT 819 a 897 c 854 g 721 t			
ORIGIN			

Query Match 53.0%; Score 1628.6; DB 10; Length 3291;
Best Local Similarity 75.7%; Pred. No. 9.7e-288;
Matches 2088: Conservative 0; Mismatches 629; Indels 42;


```
Db 2195 TCCTGTGAGACATCCGAGCATGGACCTTGAACATATGGAGGAGACCTTCCACAGAT 2254
Qy 2302 gacaagtgtggtccagagagcgtggagaatgggagactttgtcactgtgaggagctc 2361
Db 2255 GACAAGTGTCTCTCCAGAGAGGTGGCAATGGAACCTTTTGGACACTGTGAAGAGTCT 2314
Qy 2362 gggagcgcgtgtggtgtattctctccgcgcagcgggtatgagctccaagcccgaggagc 2421
Db 2315 GGAAGCTGGTGTGTACTCTCTTCCATGGATACAAAGCTGCAGGCGCAGGAGCAG 2374
Qy 2422 ctcaactgcaccaggaagatggtatttccagctccctctgcaaaagatggaacgag 2481
Db 2375 GTACATGTATCCAGAGGATGGGACTCGGAACCTCTCTGTAAAGATGTAATGAT 2434
Qy 2482 tgtgcagacggtgcccacccccctccagcctctgcaggtgcagaaacacacaaagc 2541
Db 2435 TGTGCAGATCTGACACACCCACCTTGGCACCCCTCTGCACAGTGCAGAGACCCAGGGA 2494
Qy 2542 ggttccagtgctctgcgcggaccctcactcagagttaggagacatgggaaacctgcgta 2601
Db 2495 AGCTTCAGTGTGTGCACAGACCCCTATGTCTAGGTGAGGATGAGAAGACCTGCATA 2554
Qy 2602 gactccggagagctcctcgggtgaattgagatctcatctgcgtgctctgctgctc 2661
Db 2555 GATCTGGCAGGCTACCTCGGGCATCTTGGTCTCCATTCGACTGGTGCACCTTCTCAT 2614
Qy 2662 gagggttcgcagctctcagcgtgattgaggtggagacgactgcgcactcaaa 2721
Db 2615 GATCTTCAGTGTGTGCACAGACCCCTATGTCTAGGTGAGGATGAGAAGACCTGCATA 2674
Qy 2722 tcacactccactctcgcagacagcggaggaaacctcccgagctgagatgcggaagcac 2781
Db 2675 GCCACATTCGGCATCACAGAGA-----GAGTGACACCCAGTCAGGATGCAGAAAGAT 2728
Qy 2782 caggccgtgaggacactccgcagcggcgagctcaggaactcagagcaggagagtcg 2840
Db 2729 CAGGGAGGGGATTTACACACAAAGCGCGCAGCTCAAGACACTGGACAGGAACCTGC 2787

RESULT 11
RATTPOFR 2777 bp mRNA ROD 27-APR-1993
LOCUS Rat thyroid peroxidase (TPO) mRNA, 3' end.
DEFINITION M31655
ACCESSION M31655
VERSION M31655.1 GI:207434
KEYWORDS thyroid peroxidase.
SOURCE Rat thyroid cell line FRTL-5, cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 2777)
Isozaki,O., Kohn,L.D., Kozak,C.A. and Kimura,S.
Thyroid peroxidase: Rat cDNA sequence, chromosomal localization in
mouse, and regulation of gene expression by comparison to
thyroglobulin in rat FRTL-5 cells
Mol. Endocrinol. 3, 1681-1692 (1989)
90114171
Location/Qualifiers
1..2777
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
<1..2313
/feature="thyroid peroxidase"
/codon_start=1
/protein_id="AAA42265.1"
/db_xref="GI:207435"
/translation="IRPTIGVCNRRDHPHAGNSALAKRWLPVYEDGESQPRGNWPN
FLYHGLGPSVREVRHIIQVSNVAVTDDQYDFLPVWQAVIDHIALTPGSTAA
FWGVDYCLTQENQPCFP IQLPNSRRITACLPYRSSACGTGQDQALFNLSAAN
PROQMNGITSLDASTVYSSPGVEKQLRNWSSAGLLRVNTLHLDGRLYLPFASAA
CAPEPAPHANTPCFLAGDGRASVEPALAAVHTLWLRHNRLATAFKAINTHWSANT
AYQEARVVVGLHQLIITWRDITPKILGPDPAFRQYVGPYEGYNPTVNPVSNVFNSTAA
F
```

```
RFGHATVHPLVRRINTDFQDHTELPRLQLHDVFFRPWRLIQBGGLDPIVIRGLLAPAK
LQVQQLMNEELERLELVNVTGLDASLNQGRDHLPGYENWREFCGLSRLETP
AEINKAIANRSHVNTIMELYKADNIDVWLGLGLAEKFLPGARTGFLFKQKMAL
EDGRFPWNSHVTDQROLEKHSIPRVICDNTGLTRVPVDAFRICKFKYQDPSCE
RIGSMDLRLARETFPODDKCVFPEKVDNGFVHCEESGKLVLYVSCFNGIKLQGOEV
TCQNGWDSPEPPVCKDNECADLTHPPCHSSAKKNKYGSCQCTDPIYMLIGEDSKTC
ISGRILPRASVWSIALGLLIGLGLASISWTIVICRWTHADKKNKSLILITERVTMESGRK
SDSGISQPAEYQDAEQEPAYGSRVLICE*
BASE COUNT 690 a 752 c 722 g 613 t
ORIGIN
Query Match 46.9%; Score 1441.8; DB 10; Length 2777;
Best Local Similarity 77.6%; Pred. No. 1.3e-253;
Matches 1791; Conservative 0; Mismatches 492; Indels 24; Gaps 3;
Qy 535 tacaggcccatcacaggagcttgcaaacagagagaccacccagatggggcgctcccaac 594
Db 1 TACCGGCCCATCACAGGGGTGTGCAACACAGAGATCACCCAGATGGGAGCGCTCCAC 60
Qy 595 acggccttgccagatgctccctccagctctatgagcggcttcagtcagccccgagcc 654
Db 61 ACAGCCCTAGCAGATGGCTGCTCTCTACGAAGATGGTTTCAGTCAGCCAGAGAGC 120
Qy 655 tggaaacccggctcttgtaacacgggttcccaactgccccgggtccggaggtgacaaga 714
Db 121 TGAACCCCTAATTTCTTATACCATGGCTTGGGACTGCCGTACGGAGGTGACAAGA 180
Qy 715 catgtcattcaagttaaataagattgttcacagatgattgaccgctattctgacctctg 774
Db 181 CACCTCATTTCAAGTTTCAATGAGGCTGTGCTGAGTGAAGATGACCACTACTCTGATTCT 240
Qy 775 atgcatggggacaatacatcgacacacacatcggttccacacacagagacacacaaa 834
Db 241 CCGGTGTGGGCACAGTACATCATCATCATCATCATCATCATCATCATCATCATCATCAT 300
Qy 835 gctgccttcggggaggggtctgactgacagatgactgtgagaaacaaacacatgtttt 894
Db 301 GCAGCTTCTGGGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 895 cccatacaactcccgaggaggagggcgccggcgccggcgccgctgtgtgccccctctac 954
Db 361 CCCATACAGCTTCCCTCAAACTCCTCACGGAC-----CACTGCATGCTACTCTTCTAC 414
Qy 955 cgtctctggcgcgtcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1014
Db 415 CGCTCCTCAGCGCCCTGTGGCACTGGGGACCAAGAGGACACTCTTTGGCAACCTGCTGCA 474
Qy 1015 gccaacccgagcagcagatgaacgggttgacctgttctctggagcgcgtccaccgtgtat 1074
Db 475 GCCAATCGAGGCGACAGATGAATGGCTTGACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 534
Qy 1075 ggcagctccccggcccttagagagcagctgcggaaactggaccagtcgcgaagggtgcctc 1134
Db 535 GGCAGCTCCCCAGGTGTGTGAGAAGCAGTGTGCAACTGGAGCAGCTCTGCGAGGCTGCTG 594
Qy 1135 cgcgtccacggccgcctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1194
Db 595 CGTGTCAACACCTCCACACTAGACTCTGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 642
Qy 1195 gcgcctcgcgcctgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1254
Db 643 GCATCAGCGCGCTGTGTCTCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 702
Qy 1255 ctggccgcgagacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1314
Db 703 CTGCTGGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 762
Qy 1315 ctgcgcgagcaaacccctgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1374
Db 763 CTGCGCGAGCAACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822
Qy 1375 gacgcgcgtgtgtaccaggagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1434
```


Db	721	GCAGCTCAGGACTCGGAGCAGGAGAGTGTGGATGTGGAAAGCCGGGATFACTCACAGGCTG	780			
Qy	2872	ccgagagccctctgagggcgaagtgcagagacactgcagaaacagcttctcatgttcccacaaa	2931			
Db	781	CCGAGAGCCCTCTGAGGCAAAAGTGCAGGACACTGCAGAACAGCTTCATGTTCCCAAAA	840			
Qy	2932	tcaccgtacgactcttcttccaaacacaggcaaatcggaaatcagcaggacgactgttttc	2991			
Db	841	TCACCGTACGACTCTTTTTCACAAACAGGCAAAATCGGAAATCAGCAGGACGACTGTTTTC	900			
Qy	2992	ccaaacacgggtaaatctagtaccatctcttaattactctcaggcatgataataaatgt	3051			
Db	901	CCAACACGGGTAATCTAGTACCATGTCGTAGTGTACTTCTCAGGCATGGATGAATAAATGT	960			
Qy	3052	tatatgtgca	3061			
Db	961	TATAGCTGCA	970			
RESULT 13						
HUMTP007						
LOCUS	HUMTP007	1996 bp	DNA PRI 25-APR-1996			
DEFINITION	Human thyroid peroxidase (TPO) gene, exon 8.					
ACCESSION	M25707 J02856					
VERSION	M25707.1 GI:339855					
KEYWORDS	thyroid peroxidase.					
SEGMENT	7 of 15					
SOURCE	Human lymphocyte DNA.					
ORGANISM	Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomom						
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
REFERENCE	1 (bases 1 to 1996)					
AUTHORS	Kimura,S., Hong,Y.S., Kotani,T., Ohtaki,S. and Kikkawa,F.					
TITLE	Structure of the human thyroid peroxidase gene: comparison and relationship to the human myeloperoxidase gene					
JOURNAL	biochemistry 28 (10), 4481-4489 (1989)					
MEDLINE	89352509					
COMMENT	Draft entry and computer-readable sequence for [1] kindly submit by S.Kimura,23-JUN-1989.					
FEATURES						
source						
1..1996						
/organism="Homo sapiens"						
/db_xref="taxon:9606"						
/chromosomes="2"						
/map="2pter-p24"						
<1..680						
/gene="TPO"						
/note="thyroid peroxidase Intron G"						
prim_transcript						
<1..>1996						
/gene="TPO"						
/note="thyroid peroxidase mRNA and intron"						
exon						
681..1199						
/gene="TPO"						
/note="thyroid peroxidase, (EC 1.11.1.7); G00-120-4486"						
/number=8						
1200..>1996						
/gene="TPO"						
/note="thyroid peroxidase intron H"						
BASE COUNT						
377 a 599 c 645 g 375 t						
ORIGIN						
Query Match 16.8%; Score 516; DB 9; Length 1996;						
Best Local Similarity 98.1%; Pred. No.1.6e-84;						
Matches 522; Conservative 0; Mismatches 10; Indels 0; Gaps						
Qy	893	ttcccatacaactcccgagagagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgccttctgtcgccttct	952			
Db	670	TTTCCCTGCAGCTCCCGAGGAGGCCCGCCGCGCGGCGACCGCCTGTCTGCCCTTCT	729			
Qy	953	accgcctcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgccttcttgggaactgtctcca	1012			
Db	730	ACCGCTCTTCGCGCGCTCGGCACCGGGAGCAGAGGCGCGCTCTTTGGGAACCTGTCCA	789			

LOCUS	SSTPOR1	1142 bp	mrna	MAM	13-JUL-1995
DEFINITION	Porcine mRNA for thyroid peroxidase C-terminal region (TPO).				
ACCESSION	X04645	J03463			
VERSION	X04645.1	GI:2141			
KEYWORDS	thyroid peroxidase.				
SOURCE	pig.				
ORGANISM	Sus scrofa				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
REFERENCE	1 (bases 1 to 1142)				
AUTHORS	Magnusson, R.P., Gestautas, J., Seto, P., Taurog, A. and Rapoport, B.				
TITLE	Isolation and characterization of a cDNA clone for porcine thyroid peroxidase				
JOURNAL	FEBS Lett. 208 (2), 391-396 (1986)				
MEDLINE	87054611				
COMMENT	Data kindly reviewed (23-MAR-1987) by R.P. Magnusson.				
FEATURES	Location/Qualifiers				
source	1..1142				
	/organism="Sus scrofa"				
	/db_xref="taxon:9823"				
	/cell_type="thyroid cells"				
	<1..999				
	/codon_start=1				
	/product="thyroid peroxidase (332 AA)"				
	/protein_id="CAA28306.1"				
	/db_xref="GI:2142"				
	/db_xref="SWISS-PROT:P09933"				
	/translation="FCGLSRLETWADLSAATANGRVADRLIGLHPNDIDVWLGGLA				
	ESFLPCARTGTPACILIGKQMRALRDGRFENWPGVFTAEQRRRLSRMSRVICDN				
	SGLSHVPDLDAFRVQWQPEFPCASIQGMDLGAWRAPPSGDAGCFDPVDDGDFLLC				
	EEGRQVLVFCRHGRFLRPAQITCTPRGWSPPPLCKDINCEDETDPPCHASARC				
	KNTKGGVLCECSDPLVLGEDGRTCDVAGRLPRASVSIAGLAVLGVGLAGLAWTVVCR				
	WTHADARPLLPGVEGEGDGKPSLPLPGCGRRRDVGAAPALEVEQDLSCGSRGLCE"				
	1123..1128				
polyA_signal	/note="pot. polyA signal"				
	1142				
polyA_site	/note="polyA site"				
BASE COUNT	177 a 363 c 408 g 194 t				
ORIGIN					
Query Match	16.6%; Score 510.4; DB 4; Length 1142;				
Best Local Similarity	73.8%; Pred. No. 2e-83;				
Matches	649; Conservative 0; Mismatches 231; Indels 0; Gaps 0;				
QY	1873 ttctgcggctgctgcctgcctggagagaccccgctgaacctgagacagcagcagcagg 1932				
Db	1..TTCTGCGCGCTGTCCAGACTCGAGACCTGGCGCGACCTGCTGCTCACTGCCAACCGG 60				
QY	1933 agcgtgcccagaagatccttgactgtacaagcactcctgacacacatcgatgtcgtgctg 1992				
Db	61 CGTGTGCCACACAGATCTCGGGCTGTACAGCATCCGGATACATATGACGTCTGGCTG 120				
QY	1993 ggaggcttagctgaaacttctcccgagggtcggagacgggcccctgtttgcctgtctc 2052				
Db	121 GCGGCTTGCGGAGAGCTTCTCCCTGGGCGGGACCGCGCGCTGTCGCTGCATC 180				
QY	2053 attgggaagcagatgaaggctctcggggacggtgactggttttggtggggaacagcac 2112				
Db	181 ATCGGAAAGCAGATGAGGGCCCTGAGGACCGCGGCTTCTGGTGGGAGAACCCCGGG 240				
QY	2113 gtcttcacgatcacagagcgtgactgagagaagcacctccctgtctgggtcatctgt 2172				
Db	241 GTGTTACAGAAGCCAGAGCGCGGCTGAGCGGCGGCTATGTCGCCGCTCATCTGC 300				
QY	2173 gacaacactggctcaccaggggtgcccatggatgccttcccaagtccggaattccccgaa 2232				
Db	301 GACAACAGCGGCTGTCCACAGTGCCTTTGATGCCCTTCGGGTGGGCCAGTGGSCCTCAG 360				
QY	2233 gactttgagtcttgtgacagcactcactggcagcactgagcctgaggggaacaccttt 2292				
Db	361 GAGTTCGAGCCGTGTGCAGCATCCAGGGCATCCAGGGCATGACCTGGCGCGTGGAGGAGGCCCT 420				

